



Cloning and Characterization of Glyceraldehyde-3-Phosphate Dehydrogenase from Orchid (*Cymbidium goeringii*)[†]

YUN-FANG TIAN¹, HAN LI², XIU-YUN YUAN³, YU-ZHEN YANG¹ and BO CUI^{3,*}

¹Life Science Department, Zhengzhou Normal University, Zhengzhou 450044, P.R. China

²Bioengineering Department of Puyang Vocational and Technical College, Puyang 457000, P.R. China

³Institute of Biotechnology, Zhengzhou Normal University, Zhengzhou 450044, P.R. China

*Corresponding author: Tel: +86 371 65502139; E-mail: cuib2014@163.com

AJC-15690

Total RNA of *Cymbidium goeringii* (*C. goeringii*) leaves was extracted with trizol. With primers designed based on the sequence of glyceraldehyde-3-phosphate dehydrogenase (GAPDH) from NCBI, a full-length cDNA of GAPDH was cloned from *C. goeringii* by RT-PCR combined with RACE. The newly cloned GAPDH gene was designated as CgGAPDH1 (JQ978220). Sequence analysis shows that CgGAPDH1 is 1 424 bp in length, contains a single open reading frame from 168-1190 bp, encodes proteins with 340 amino acids. CgGAPDH1 showed 99 % similarity with GAPDH5 (JN177726). Phylogenetic analysis based on the sequence of GAPDH showed that *C. goeringii* had closely relation with *Triticum aestivum* and *Lilium longiflor*. Bioinformatics analysis indicated that CgGAPDH1 protein is soluble without transmembrane domain. Homology modeling of the GAPDH in *C. goeringii* based on the 3D structure is similar to that of homologues in *Triticum aestivum*, *Oryza sativa* and *Zea mays*, which firmed that the polypeptides were highly conserved.

Keywords: Glyceraldehyde-3-phosphate dehydrogenase, Cloning, Bioinformatics, RACE.

INTRODUCTION

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) played an important role in molecular biology research^{1,2}. GAPDH is one of the most commonly housekeeping genes used in expression studies³⁻⁶. GAPDH also has many unknown roles appeared in anaerobic stress, heat shock, sucrose supply, wound-stress, *etc.*⁷⁻¹⁰. Orchidaceae is the largest family of flowering plants, with at least 24000 species¹¹. Molecular biology of orchid gene has not been widely performed, therefore, effective molecular breeding of this plant is not available. In the genus *Cymbidium goeringii* (*C. goeringii*), no information about GAPDH is currently available. The novel GAPDH in *C. goeringii*, designated as CgGAPDH1, was isolated and analyzed using bioinformatic software. All of these would provide the ground work for understanding the function of CgGAPDH1 as house keeping gene.

EXPERIMENTAL

During the period of vegetative and reproductive, roots, leaves, tepals, ovaries and pedicels were sampled from *C. goeringii* and immediately frozen in liquid nitrogen and stored at 80 °C.

RNA isolation and analysis: Total RNA was extracted from materials using trizol. Frozen plant material was pulverized with mortar and pestle in liquid N₂ 100 mg of powdered tissue was used to isolate total RNA. The RNA was precipitated and the pellet was dissolved in diethylpyrocarbonate (DEPC)-treated H₂O. The integrity of RNA samples was assessed by agarose gel (1.2 %) electrophoresis. Concentration and purity of RNA were determined from the A260/A280 ratio using a UC800 nucleic acid-protein analyzer.

Reverse transcriptase (RT)-PCR: For the isolation of GAPDH cDNA fragments, 1 µg of total RNA was used to synthesize first strand cDNA from the polyA tail with a T(18) primer using the cDNA-for-PCR kit (TaKaRa Biotechnology, Dalian, China).

Isolation of GAPDH of *C. goeringii*: Degenerate primers (5' primer: 5'-AA(G/A) ATCGGA ATCAA (T/C)GG(G/A)TTCG-3' and 3' primer: 5'-AT(G/A) CT(G/T) GACCTG(C/T)TGTCACC-3') were designed based on the sequences of GAPDH in GeneBank. RT-PCR was performed using Taq DNA polymerase (TaKaRa Biotechnology, Dalian, China). The PCR program was as follows: 5 min denaturation at 94 °C, followed by 35 cycles of a denaturation step at 94 °C, an annealing step

[†]Presented at 2014 Global Conference on Polymer and Composite Materials (PCM2014) held on 27-29 May 2014, Ningbo, P.R. China

at 56 °C and an extension step at 72 °C. Each step was 45 s long and the final extension step was at 72 °C for 10 min. The PCR products of amplification were separated using 1 % agarose gels and stained with ethidium bromide. The PCR fragment of the predicted size (850-bp) was excised from the agarose gel and purified using Montages DNA Gel Extraction Kit (Tiangen, Beijing, China). The purified PCR products were ligated into a pMD19-T vector (TaKaRa Biotechnology, Dalian, China) for identification and sequencing.

Cloning of GAPDH cDNAs full-length from *C. goeringii*: The full-length cDNA of GAPDH was isolated using RACE (rapid-amplification of cDNA ends). According to conserved regions of the sequence of GAPDH, primers (5' race: outer-5'- AACATAGTCAGCACCAGCCTCAC-3' and inner-5'- GTATTCGGTGGTGATGAACGGAT-3' 3' race: outer-5'- ATGACAGGTTTGGCATTGTTGAG-3' and inner-5'- AAGGCTGTTGGTAAA GTTCTCCC-3') were designed and synthesized. The RACE PCR was performed using the Clontech Smart Race kit.

The PCR products were separated by electrophoresis on a 1 % (w/v) ethidium bromide-stained agarose gel. The PCR fragment of the predicted size (300 and 650 bp) was excised from the agarose gel and purified using Montages DNA Gel Extraction Kit (Tiangen, Beijing, China). The purified PCR products were ligated into a pMD19-T vector (TaKaRa Biotechnology, Dalian, China) for identification and sequencing. Both 300 and 650 bp products were aligned and assembled to one sequence. The assembled sequence was re-amplified. The sequence was submitted to Genbank nucleotide databases after being sequenced and verified. The ORF was determined by the program of NCBI ORF Finder.

Sequence analysis of CgGAPDH1: The obtained sequence CgGAPDH1 was analyzed using bioinformatics tools. The sequence comparison was conducted through database search using BLAST program. The phylogenetic analysis of GAPDH proteins was aligned with DNAMAN using default parameters. Physicochemical properties of protein sequence were synthetically analyzed through ProtParam. Trans-membrane domain was predicted on-line using TMHMM. The hydrophilicity/ hydrophobicity of CgGAPDH1 protein was analyzed with the help of ProtScale. Subcellular localization of CgGAPDH1 was identified by using TargetP and SubLoc. Structure of CgGAPDH1 was predicted by Sopma and Predict-protein. The protein structure was modeled on Swiss-model server.

RESULTS AND DISCUSSION

Cloning of GAPDH cDNAs fragment from *C. goeringii*:

Analytical results of the extracted total RNA showed that the ratio of A_{260} to A_{280} was 1.98 and the 18S and 28S rRNA bands (Fig. 1) were specific, indicating that the total RNA has high purity and is not decomposed.

The homologous GAPDH cDNAs designated CgGAPDH1 (JQ978220), was isolated from *C. goeringii*. The full-length cDNA sequence of GAPDH is 1 424 bp in length (Fig. 2), containing a 167-bp 5' non-coding region, a 234-bp 3' non-coding region, a 1 023 bp ORF, encodes for a 340 aa predicted polypeptide. The sequences of the CgGAPDH1 cDNAs were

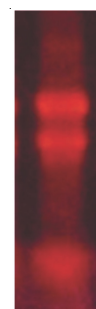


Fig. 1. RNA extraction

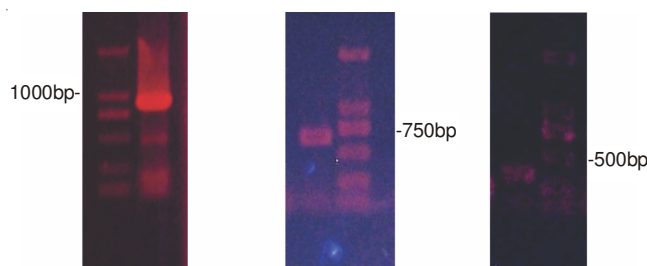


Fig. 2. A. RT-PCR products; B. 3'RACE-PCR products; C. 5'RACE-PCR products

more homologous among monocots and CgGAPDH1 shared 99 % similarity with *Cymbidium faberi* GAPDH5 (JN177726).

Amino acid sequence comparisons, phylogenetic analysis:

Published GAPDH genes that showed a high degree of homology to the CgGAPDH1 were selected for the multiple alignment process (Fig. 3). In this comparison the CgGAPDH1 present highly similarity to the known GAPDH proteins, especially *Triticum aestivum* and *Lilium longiflorum*.

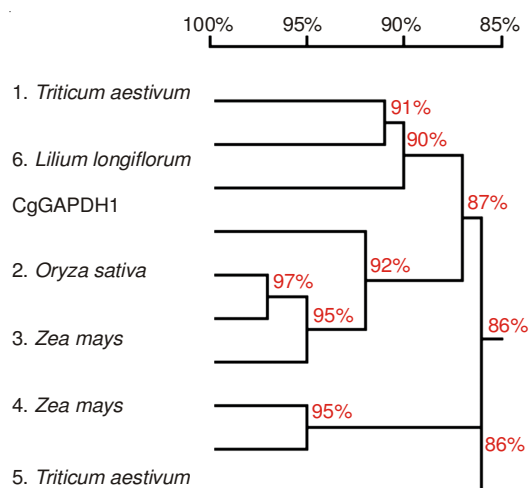


Fig. 3. Phylogenetic tree of CgGAPDH1 protein sequences with known GAPDH protein sequences

Physico-chemical properties of CgGAPDH1 protein analysis:

CgGAPDH1 contains an ATG start codon and TAA stop codon. The deduced polypeptide has 340 aa with a predicted molecular size of 36832.0 and pI at 6.27. The peptide has 2 cysteine and relatively higher content of Val (10.9 %), Ala (10.0 %), Gly (8.5 %) and Lys (8.5 %), each above 8 % (mol % total amino acid). Total number of atoms is 5215, formula is $C_{1652}H_{2618}N_{442}O_{496}S_7$. Among them, the number of

residues with negative charge (Asp+Glu) was 44,13.0 % of the total amino acid amount; the number of the residues with positive charge (Arg + Lys) was 41, 12.0 % of the total amino acids. The N-terminal of the sequence considered is M (Met). The estimated half-life is 30 h. The instability index (II) is computed to be 23.04 and the protein is stable. Grand average of hydropathicity (GRAVY) is -0.080.

Characteristics of CgGAPDH1 in hydrophilicity, hydrophobicity, trans-membrane domain and subcellular localization: TMHMM analysis indicated that CgGAPDH1 protein did not have obvious transmembrane domain. The hydrophilicity/hydrophobicity analysis result showed that the hydrophilicity/hydrophobicity of CgGAPDH1 protein were 2.011 in maximum at position 24, -2.433 in minimum at site 324. The number of hydrophilic amino acid residues is more than hydrophobicity amino acid residues in one peptide chain. Therefore, the putative CgGAPDH1 protein were hydrophilic protein.

The prediction of subcellular localization of CgGAPDH1 with TargetP and SubLoc showed that CgGAPDH1 protein likely belongs to secretory pathway signal peptide and cytoplasmic.

Prediction of secondary structure and function domain of CgGAPDH1: The prediction results of secondary structure of CgGAPDH1 by Sopma showed that the secondary structure of CgGAPDH1 consisted of random coils (35.00 %), α -helices (32.65 %), extended strands (23.53 %) and β -turns (8.82 %). The secondary structure proportion of random coil and α -helix was higher. The prediction results by predictprotein showed that the secondary structure of CgGAPDH1 consisted of loops (41.18 %), helices (27.06 %) and strands (31.76 %), which is identical to that of Sopma.

Three-dimensional modelling of the proteins analysis: The final step in the bioinformatical analysis of the protein subunit structures was their 3D modeling by means of the SWISS MODEL program, which predicts the protein structure based on amino acid sequences and comparisons with structures contained in databases. The 3D views of the resulting models (Fig. 4) were generated on-line modeling. The 3D model (Fig. 4) is the visual representation of the information contained in the CgGAPDH1, which showed spatial arrangement of secondary structure elements like α -helices, β -strands, loops and turns. Comparing with 3D structures of *Triticum aestivum*, *Oryza sativa*, *Zea mays*, the structure of CgGAPDH1 is conserved.

In this study, with the intention of cloning GAPDH from *C. goeringii*, the combined strategy using RT-PCR and RACE-PCR was performed to isolate full-length cDNA CgGAPDH1. For an orchid plant like *C. goeringii*, insight into internal control gene is of great interest. Knowledge of genes may lead to reasonable experimental protocols for relative quantitation of gene expression. The isolated gene in the present study contained two conserved domains. Multiple sequence comparison and the phylogenetic analysis suggests that the isolated GAPDH in *C. goeringii* be most likely homologues with other monocot.

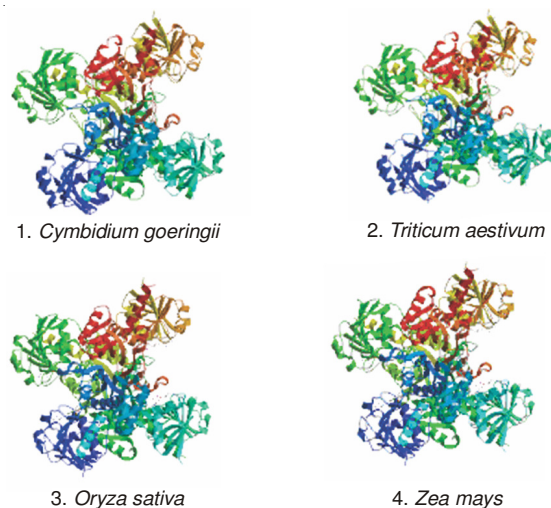


Fig. 4. GAPDH protein spatial structure predicted by Swiss-Model

Characters were analyzed on-line using bioinformatics. However, some study results are needed further to confirm. For example, bioinformatical 3D modeling only suggests possible protein structures that have to be verified experimentally by appropriate techniques such as X-ray crystallography and/or NMR spectroscopy. The information provided by the presumptively folded states of the proteins is indispensable for further directed study and understanding of their function.

GAPDH mRNA expression level in different organs and growth periods will reveal that whether differences is significant and the abundance is higher and CgGAPDH1 could be an internal control gene of *C. goeringii*. However, whether CgGAPDH1 is the most appropriate internal control gene of *C. goeringii*, which need to be further studied.

ACKNOWLEDGEMENTS

This work was supported by the Scientific and Technological project of Henan Province(092102110128) and the Education Department of Henan Province Science and Technology Research projects (14A220005). The authors thank Prof. Bo Cui (Institute of Biotechnology, Zhengzhou Normal University) for his collaboration in the treatment of the plant materials.

REFERENCES

1. E.J. Tisdale, *J. Biol. Chem.*, **277**, 3334 (2002).
2. M.A. Sirover, *Protein Struct. Molecul. Enzymol.*, **1432**, 159 (1999).
3. K. Dheda, J.F. Huggett, S.A. Bustin, M.A. Johnson, G. Rook and A. Zumla, *Biotechniques*, **37**, 112 (2004).
4. H.B. Yan, W.Y. Ge, Y.D. Cheng, J.-G. He and J.-F. Guan, *J. Northwest Sci-Tech Univ. Agric. Forestry*, **40**, 181 (2012).
5. C. Luo, X. He, Y. Hu, C. Tan and S.-J. Ou, *J. Fruit Sci.*, **28**, 1019 (2011).
6. G.F. Zhang, S.H. Jiang, B. Cui, X.-Y. Yuan, Y.-F. Tian and Y.-Z. Ye, *J. Henan Agric. Univ.*, **47**, 68 (2013).
7. M.A. Sirover, *J. Cell. Biochem.*, **66**, 133 (1997).
8. M.A. Sirover, *J. Cell. Biochem.*, **113**, 2193 (2012).
9. Y.F. Tian, S.H. Jiang, X.Y. Yuan and B. Chui, *Acta Agric. Univ. Jiangxiensis*, **35**, 870 (2013).
10. G.L. Fu and X.H. Huang, *Acta Biophys. Sin.*, **29**, 181 (2013).
11. M.F. Fay and M.W. Chase, *Ann. Bot.*, **104**, 359 (2009).